Auditors' Error Handling, Error Orientation, and Audit Organizations' Error Climate — Validity and Reliability of a Measurement Instrument

Ulfert Gronewold
University of Potsdam

Michaela Donle
Stiftung Warentest

ABSTRACT: The auditor's behavior may have a considerable impact on the results of an audit and the auditee's acceptance of the auditor's recommendations. A variable of particular importance is the auditor's error handling behavior, i.e. the auditor's behavior after detecting an auditee's error. We developed a questionnaire for measuring auditors' error handling behavior and also for two related constructs: audit organizations' error climate and auditors' individual error orientation.

In this study, we empirically investigate the questionnaire's validity and reliability. We use confirmatory factor analysis (CFA), Cronbach Alpha, and the Fornell-Larcker-criterion on survey data from 284 external, internal, and public sector auditors. Results are promising – the measurement instruments for error climate and error orientation, which were adapted from questionnaires that have been formerly tested in psychological research, achieved very good values. The (originally developed) measurement instrument for error handling achieved rather moderate but still satisfactory values in the analyses.

Key Words: Error Handling, Organizational Error Climate, Individual Error Orientation, Confirmatory Factor Analysis.

INTRODUCTION

The auditor's behavior during an audit may have an important impact on the audit's results. In particular, the auditor's behavior following the detection of an auditee's error may have important consequences for the acceptance of his[1] findings and the implementation of his recommendations. As an audit's central benefit does not only result from detecting errors, but also from improving structures and processes (Richter 2000, 665), acceptance of the auditor's recommendations by the auditee is essential for audit effectiveness. Under favorable conditions, a good auditor-auditee-relationship may facilitate a more rapid and straightforward achievement of the audit goals, so that audit efficiency might also improve. A crucial question is how auditors deal with detected errors. Generally speaking, errors may be considered either as a result of personal failure that must be punished, or as an occasion for learning (Richter 1999, 355).

Due to its relevance, audit practice has called for more research in this field. A basic demand concerns research on the auditor's relationship to the auditee – presumably with the objective to achieve improvements and more benefits from this relationship for both sides. Indirectly, this claim also refers to a better handling of the auditee's errors by the auditor (e.g. Haakenson 1976, 32-40; Baker 1978, 58-62; Sheffield 1982, 32-33; Campfield 1983, 49-52; Keating 1995, 59-61; Allen 1996, 58-61; Kimbrough 1997, 64-66; Burke and Kovar 1999, 11-15; Jeffords et al. 2001, 17-24; Chadwick 2002, 52-55; Ridley 2003, 34-36). Better reporting and communication by the auditor are also claimed (e.g. Holman 1981, 39-44; Maniak 1990; Crockett 1992; Whitham 1996, 60-62; Didis 1997, 36-38; Schwartz 1999, 48-49; Lynch and Golen 2003, 53-57; Wells 2003, 75-78; Williams 2003, 50-57). However, to date, behavioral and experimental auditing research has

[1] For reasons of better readability, the masculine form will be used throughout the paper. It equally represents the feminine and masculine form.

dealt with the interaction behavior between the auditor and the auditee to a far smaller extent than it has dealt with auditors' cognitive processes, judgment, and decision making. An exception is a study concerning auditors' styles carried out by the Institute of Internal Auditors (1972), which showed that the auditor's behavior directly influences the performance of the audited department.

The auditor's behavior after detecting an auditee's error as a specific and particularly relevant aspect of the interaction behavior between auditor and auditee has not yet been studied in auditing research. However, findings in industrial and organizational psychology suggest that individuals differ in their error handling behavior. This research also showed that different error handling strategies have different impacts on performance and learning for both individuals and organizations (Edmondson 1996,; Göbel and Frese 1999, 101).

We developed a questionnaire for measuring auditors' error handling behavior, auditors' error orientation, and audit organizations' error climate. Error orientation is the auditor's attitude towards his own errors. Error climate refers to the manner in which auditors' errors are typically dealt with in the auditor's own organization. Thus, these two constructs refer to auditors' *own* errors, whereas the error handling construct refers to the auditor's way of dealing with the *auditee's* errors. While error orientation, error climate, and error handling behavior might be interrelated to a certain degree, they represent three distinct constructs.

The main purpose of this study is to provide evidence on the validity and reliability of the questionnaire and its scales and items that were developed for measuring the three constructs. Information on the quality of new measurement instruments is important for future empirical research that will rely on them. Our investigation is based on a survey with German auditors who filled out the questionnaire (overall 284 usable responses). The Validity and reliability assessments reported in this study rely on measures provided by confirmatory factor analysis

(CFA)[2], Cronbach Alpha (Cronbach 1951), and the Fornell-Larcker-criterion (Fornell and Larcker 1981, 45-46).

The remainder of this paper is organized as follows. The next section deals with the theoretic background. After defining central terms, a conceptual framework including the constructs investigated in this study (error handling, error climate, and error orientation) is described, followed by theoretic descriptions of each of these constructs. In the following main section, concepts of validity and reliability are outlined first. Then, the development of the questionnaire, the subjects of our study, the method of data collection, and the method of analysis are described. The empirical results are presented in the next main section. The concluding main section, the discussion, includes a summary and evaluation of the main findings, limitations of the study, and issues for future research.

THEORETICAL BACKGROUND

Definition of Terms: “Error” and “Error Handling”

For auditors, an important task is to detect errors or deviances in the audited issues. In auditing research, the term “error” is normally defined as divergence between a recorded value and the correct or “true” value (e.g. Baetge 1970, 33; Egner 1980, 180; Von Wysocki 1988, 140; Buchner 1997, 234; Kinney 2000, 93; Lindner 2003, 411). There are different classifications of errors. Usually, a difference is made between intentional and unintentional errors (Mautz and Sharaf 1961, 120; Wanik 1985, 853; Von Wysocki 1988, 249; Buchner 1997, 241), as well as between material and immaterial errors (Mautz and Sharaf 1961, 118; Schulte 1970, 22; Egner 1980, 185; Von Wysocki 1988, 248-249). In this study, the term “error” refers to the auditor’s

[2] All CFA reported in this study were conducted by means of the software package AMOS 5.

perception of any divergence of an actual or "real" status from an ideal or "true" status, which is detected by the auditor during the audit. Intentional errors are not considered because the behavior facing fraud presumably differs significantly from the behavior facing unintentional mistakes. Aside from infringements of rules, inefficiencies shall be included in the error concept as well. It should be stressed that the error concept refers to the *auditor's perception* of an error and does not differentiate between objective and subjective errors. The focus is only on what the auditor himself considers to be an error and how he would deal with it.

Following the psychological term of error handling (Zapf et al. 1999, 406), in this study "error handling" refers to the auditor's behavior after detecting an auditee's error. The process of error handling includes the explanation and elimination of the error. This study focuses on the auditor's interaction behavior with the auditee, not on cognitive processes.

Conceptual Framework

Investigations of the auditor's error handling should also take into account related constructs that might be relevant. Psychological research has dealt with a person's individual error orientation and an organization's error climate. These constructs were adapted to an auditing context and included into this study. Error orientation is the auditor's personal attitude towards his own errors. Error climate refers to the manner in which auditors' errors are typically dealt with in the auditor's own organization. Thus, these two constructs refer to auditors' *own* errors, whereas the error handling construct refers to the auditor's way of dealing with the *auditee's* errors.

Figure 1 shows a conceptual framework that arranges the three constructs as belonging to the auditor's behavior (error handling), the auditor's personal characteristics (error orientation), and the situation (error climate). As personal characteristics, situational factors, and behavior in general may be interrelated to a certain degree, so may be the three aforementioned constructs, which is symbolized by the arrows in the figure. An auditor's handling of the auditee's errors may likely be influenced by the auditor's personal attitude towards own errors (error orientation) and by the error climate within his own organization. Also, as an audit organization is formed by the sum of its members, auditors' individual error orientation may affect the error climate of their organization, because this climate may be seen as the sum of the attitudes of the organization's members. Vice versa, the organizational error climate may influence an individual's error orientation if it causes some pressure of adaptation or if recruitment of new staff is done in a way that only auditors with error attitudes similar to the organization's error climate are hired. Finally, under certain circumstances, error handling behavior might also have some backward effects on error orientation and error climate, e.g. if manners of dealing with the auditee's errors are perceived to be beneficial, which therefore are adopted for dealing with own errors as well.

Despite these expected interrelationships between the three constructs in the triangular form depicted in Figure 1, it should also be clear that error handling, error orientation, and error climate represent three distinct constructs. In the following subsections each is described in more detail.

[Insert Figure 1 about here]

Strategies of Error Handling Behavior: Error Prevention and Error Management

Two typified strategies of auditors' error handling behavior may be distinguished: error prevention and error management (Frese 1995, 112; Frese and Van Dyck 1996). These strategies become manifest both within the oral and written feedback to the auditee about detected errors. The strategy of error prevention implies that the auditor evaluates errors negatively and that the audit's main objective is to detect as many material errors as possible, document them, and get them eliminated. Future errors shall be avoided by means of deterrence. After detecting an error, an auditor who follows a strategy of error prevention concentrates on identifying the individuals responsible for the error. Recommendations are determined by the auditor himself without coordinating with the auditee. The preferred form of feedback is a written report that contains exclusively negative audit findings without including any remarks or comments by the auditee.

On the other hand, an auditor who applies a strategy of error management understands himself as a specialist who makes available his knowledge to the auditee and the audited department. The audit's primary objective is organizational development and, as far as errors have been detected, to learn from them. Future errors shall be avoided by means of learning. After detecting an error, the main focus is not to look for the responsible or "guilty" person, but to analyze systematically the error's causes. Through his work, the auditor intents to train the auditee's error competence, so that the auditee becomes able to detect and correct errors on his own and to avoid repeating the same errors in the future. Dealing openly with errors is encouraged. The preferred form of feedback is oral communication with the individuals involved in the system. The auditor is interested in the auditee's point of view and in considering his opinion when he forms his judgment. His report usually contains a thorough evaluation of the organization, including both positive and negative aspects. The auditee's observations and objections are also documented in the report.

Both strategies described above represent extreme behavioral patterns that mark the end points of a continuum. Real auditor behavior will most likely be situated in between of the end points and integrate aspects from both error prevention and error management. However, there may be a tendency towards one of the two strategies for a specific auditor in a specific situation.

Individual Error Orientation

The individual error orientation refers to a person's attitude towards his own errors and to the manner the person deals with these errors (Rybowiak et al. 1999, 527). Error orientation depends on how negatively own errors are evaluated and on the extent to which errors are prevented. It is also connected to specific activities following the error occurrence. This includes to what extent the erring person talks openly about committed errors, to what extend he analyzes errors that occurred, and, finally, to what extent he actively deals with errors and tries to learn from them (Rybowiak et al. 1999, 529). A person with an error prevention orientation generally considers errors as negative and actively tries to avoid them. If an error still occurs, he tends to conceal the error and does not take time to analyze the error's cause(s). He has only a low competence in dealing with errors and does not focus on learning from them.

In contrast, a person with an error management orientation primarily associates errors with the opportunity to learn and to improve work routines, processes, and structures. He anticipates that errors may occur even if work is done carefully. Finally, he communicates openly about his own errors and deals actively with them. Psychological research suggests that individual error orientation may be measured by means of the subscales "communication about errors", "learning from errors", "analysis of errors that occurred" and "competence in dealing with errors" (Rybowiak et al. 1999).

Organizational Error Climate

The organizational error climate is a specific aspect of the organizational climate. It comprises beliefs, attitudes, and norms concerning errors, which are shared by all members of the organization. It also includes the way errors are usually dealt with (Van Dyck 2000, 106). Similar to the individual error orientation, this concept refers to "own" errors, i.e. errors caused by members of the organization. An organization with an error climate that tends towards error prevention favors behaviors that corresponds to the error prevention strategy. This strategy may be adequate for an organization in stable conditions where it is not necessary to adapt to a changing environment (Sitkin 1992, 233; Van Dyck 2000, 36). Other organizations support an error management behavior. Prior research showed that these organizations achieved better objective and subjective performance parameters than organizations with a culture of error prevention (Edmondson 1996; Göbel and Frese 1999, 101). Psychological research also suggests that organizational error climate can be measured by means of the subscales “help when an error occurred”, “analysis of errors that occurred”, “communication about errors”, and “correction of errors” (Van Dyck 2000).

RESEARCH METHOD

Concepts of Validity and Reliability

The purpose of this study is to provide evidence on the validity and reliability of the measurement instrument, i.e. the questionnaire that was developed for measuring auditors’ error handling, error orientation, and audit organizations’ error climate.

Validity is the degree to which a measurement instrument is able to accurately measure the real concept that is intended to measure (e.g. Stier 1999, 56; Hair et al. 1998, 584). In other words, it refers to the freedom from systematic measurement errors. The best way to assess validity would be to confront the measurement instrument to be tested with an external validation criterion. However, such a criterion is not available in many cases, so that other means of validation must be applied. This is also true in the present case.

An established and appropriate way is to refer to construct validity. Generally speaking, “[c]onstruct validity assesses whether a measure relates to other observed variables in a way that is consistent with theoretically derived predictions" (Bollen 1989, 188). It comprises nomological validity, convergent validity, and discriminant validity (Homburg and Pflesser 2000, 421-422). Nomological validity refers to the degree to which predicted relationships between a construct of interest and other constructs can be confirmed, with the relationships being determined in advance on a theoretical basis (Homburg and Pflesser 2000, 422; Bagozzi 1979). Convergent validity refers to the degree to which two or more different measurements of the same construct correspond to each other (Homburg and Pflesser 2000, 421; Bagozzi and Phillips 1982, 468). Finally, discriminant validity refers to the degree to which measurements of distinct constructs differ from each other (Homburg and Pflesser 2000, 421; Bagozzi and Phillips 1982, 469).

In this study nomological and discriminant validity are assessed. For assessing nomological validity, the three constructs that are measured in the questionnaire (error handling, error climate, and error orientation) are related to each other in correspondence to the theoretic framework described above (see Figure 1). Factor models are specified and tested by CFA. CFA serves for assessing both nomological and discriminant validity, whereas discriminant validity is also assessed by means of the Fornell-Larcker-criterion.

Reliability is the degree of freedom from random measurement errors (Homburg and Pflesser 2000, 420). Usually, three different ways of empirically determining reliability are distinguished: test-retest-reliability, parallel-test-reliability, and internal-consistency-reliability (Homburg and Pflesser 2000, 421; Stier 1999, 53). In this study, internal-consistency-reliability will be examined. Internal-consistency-reliability is the degree to which different items that measure the same construct are consistent in their measurements (Hair et al. 1998, 583; Stier 1999, 55-56). It will be assessed for the specified measurement models of the three constructs that are addressed by the questionnaire by using both CFA and Cronbach Alpha.

Development of the Measurement Instrument

For measuring error handling, error climate, and error orientation, a questionnaire containing measurement instruments for the three constructs was developed. The questionnaire's first section addresses the auditor's error handling in an audit situation. The results from this section shall provide insight into the auditor's specific behavior and reveal which strategy of error handling he prefers. In the second section, the audit organization's error climate is addressed. Auditors were asked how errors are usually handled within their organization. The third section refers to the auditor's individual error orientation, i.e. the auditor's general attitude towards errors. For the second and third part (error climate and error orientation), we transferred existing measurement instruments that have been applied and tested in psychological research to an auditing context. As instruments for measuring error handling behavior were not yet available, we developed an own instrument.

In a pretest with 60 subjects (16 external, 41 internal, and 3 public sector auditors from Germany) we checked the reliability of a first version of the questionnaire. Based on the results, wordings of inappropriate items were adapted, or such items were entirely removed from the questionnaire. We obtained a revised version of the questionnaire with internal consistency scores (Cronbach Alpha) of 0.62 (error handling), 0.90 (error orientation), and 0.93 (error climate), respectively.

For measuring the auditor's error handling we use 12 originally developed items that describe an auditor's behavior in specific situations. For each item, subjects were asked how often they behave as described by the item. They provided their answers on 5-point rating scales that ranged from 1 ("never") to 5 ("regularly"). Organizational error climate and individual error orientation were both measured by means of adapted items of already existent questionnaires from psychological research. Subjects were asked to rate to what extent each item's statement applied to their own audit organization or to themselves, respectively. Subjects provided their answers on

5-point rating scales that ranged from 1 (“applies completely”) to 5 (“does not apply at all”). The end points of all scales had similar meanings with regard to error management vs. error prevention: “1” represented an error management behavior, climate, or orientation, and “5” represented an error prevention behavior, climate, or orientation, respectively.

The items of each of the three constructs were assigned to different subscales within the corresponding measurement models, which was done in accordance to the theoretic descriptions of the concepts (see above, Theoretical Background). The error-handling-items were divided into the subscales „oral feedback“ and „written feedback“ (see model specification in Figure 2), the error-climate-items were assigned to the subscales “help when an error occurred”, “analysis of errors that occurred”, “communication about errors”, and “ correction of errors” (see model specification in Figure 3), and the error-orientation-items were assigned to the subscales “communication about errors”, “learning from errors”, “analysis of errors that occurred” and “competence in dealing with errors” (see model specification in Figure 4).

[Insert Figures 2, 3, and 4 about here]

Finally, in a fourth section of the questionnaire, subjects were asked to provide some demographic data, including professional experience, age, gender, and number of auditors in the audit organization. Also in part four, we asked how frequently the auditor usually audits the same entities or departments. Finally, questions about the auditor’s job satisfaction were included.

Three versions of the questionnaire were developed, one for internal, one for external, and one for public sector auditors. However, these versions differed only marginally in wording and demographic data gathered. The final questionnaire for public sector auditors (translated to English) is included in the Appendix.

Subjects and Data Collection

The questionnaires were distributed to internal, external, and public sector auditors from Germany at conferences and training sessions. Questionnaires were also sent to auditors through professional organizations or directly by mail. Response rates varied depending on the selection method. The lowest rate was achieved for subjects contacted by chance (7 percent) and the highest rate for subjects contacted directly (95 percent). Overall, 1020 questionnaires were sent out. 284 usable responses were received (i.e., a response rate of 27.8 percent).[3]

Most of the public sector auditors who participated come from local governmental audit departments. 56 percent of the participating internal auditors work in banks or insurance companies, 17 percent in the service sector, and 14 percent in industrial enterprises. 19 percent of the internal auditors are CIAs, 36 percent are currently preparing for the CIA-exam, 10 percent plan to do so, and 34 percent do not have any plans in this regard. Most of the external auditors who participated in the study come from small audit organizations with up to 5 auditors (46 percent). 18 percent work in middle-sized companies (5 to 100 auditors) and 34 percent in big companies (more than 100 auditors). Overall, 284 auditors participated in the survey (83 external auditors, 108 internal auditors, and 93 public sector auditors). Demographic data of the sample is summarized in Table 1.

[Insert Table 1 about here]

Method of Analysis

Traditional methods for assessing the validity and reliability of measurement instruments use Cronbach Alpha, item-to-total-correlations, or exploratory factor analysis (Homburg and Giering

---

[3] Missing values were replaced by the means of the corresponding variables, which is a commonly applied technique (Hair et al. 1998, 54). Overall 122 values were replaced (i.e., only 1.1 percent of all values were affected).

1998, 118-119). However, due to various problems, these methods are subject to criticism in the literature.[4] A more powerful tool for assessing validity and reliability is CFA. As it is a theory-driven method, it is not subject to the usual criticism that affects the traditional methods (Homburg and Giering 1998, 121). CFA does not distinguish strictly between validity and reliability, i.e. it does not distinguish between systematic and random errors. However, on the one hand it offers specific measures of reliability (construct and indicator reliabilities) that allow a more differentiated assessment of internal consistency as compared to Cronbach Alpha, for example. On the other hand, by means of global fit measures validity can also be assessed (particularly nomological and discriminant validity).

Specific Measures of Reliability

In a first step, the measurement models for the three constructs (see model specifications in Figures 2, 3, and 4) are assessed. Due to its wide application, we first calculate Cronbach Alpha for the entire constructs and its subscales. As more differentiated measures of reliability, we then calculate indicator reliabilities (e.g. Homburg and Giering 1998, 124) for each subscale of the constructs by means of CFA. The indicator reliability is the proportion of the total variance of an indicator that can be explained by the underlying construct. Its value ranges between 0 and 1, with higher values indicating higher reliability. Values over 0.5 are usually regarded as acceptable (Hair et al. 1998, 612). Finally, we calculate the construct reliabilities of the three constructs and of each subscale. Construct reliability is the degree to which a construct is measured by the totality of indicators assigned to it (Homburg and Baumgartner 1998, 361). Values above 0.6 are deemed acceptable (Bagozzi and Yi 1988, 80).

---

[4] For instance, Cronbach Alpha is criticized for being positively dependent on the number of indicators and for supposing that all indicators are equally reliable (Homburg and Giering 1998, 120).

Assessment of Individual Constructs' Models

In a second step, we assess the empirical support provided by our data to the three theoretically derived measurement models for the individual constructs. This assessment relies on CFA's global fit measures. Generally, both descriptive and inferential measures may be applied. Measures of statistical inference are based on statistical tests of significance (Homburg and Pflesser 2000, 427; Homburg and Baumgartner 1998, 352-353). The best known measure of this category is the chi^2-test. However, this test has been subject to criticism, as it tests whether a model is "correct" in an absolute sense. As models are never perfect, but rather provide a more or less accurate approximation to reality, the usefulness of the chi^2-test may be actually questioned (Cudeck and Browne 1983, 164-165; Homburg and Baumgartner 1998, 353). Moreover, the chi^2-test is contingent on restrictive requirements (e.g. normal distribution of all variables), that are often not met in practice (Backhaus et al. 2006, 379-380). For these reasons, instead of applying inferential measures, we apply a set of descriptive measures[5], which are less restrictive in their requirements and better meet the practical demands. With these fit indices, model fit is assessed against experience-based values.

As an absolute fit measure we calculate the Goodness of Fit Index (GFI) (Hair et al. 1998, 655), which is considered to be the best known measure of this category (Homburg and Baumgartner 1998, 359). The general problem of these indices is that the indicated fit also improves just when the number of free parameters increases (Nunnally and Bernstein 1994, 565). Therefore we calculate three indices that take into account the degrees of freedom and thus are more precise in this respect: the CMIN/DF (Minimum Discrepancy/Degrees of Freedom), the Root Mean Square

[5] We calculate several measures in order to balance the weaknesses of individual measures as good as possible. For critiques on individual measures see e.g. Hair et al. (1998, 658-659); for critiques on using baseline-models –so that these critiques affect incremental models– see Sobel and Bohrnstedt (1985, 175).

Error of Approximation (RMSEA), and the Adjusted Goodness-of-Fit Index (AGFI) (Hair et al. 1998, 657).

Besides absolute measures of fit, we calculate incremental fit measures, which compare the specified model with a baseline model, so that these measures allow an assessment of the relative goodness of the models. Due to its widespread application, we calculate the Normed Fit Index (NFI). However, as the NFI is dependent on the sample size, we additionally calculate the Comparative Fit Index (CFI) and the Tucker-Lewis-Index (TLI), which do not have this problem (Bearden et al. 1982; Schermelleh-Engel et al. 2003, 40-41).

On the one hand, by means of these fit measures we assess for every of the three first sections of the questionnaire an unstructured model. In these models, all indicators (i.e. items) directly load on the corresponding construct (error handling, error climate, or error orientation, respectively). On the other hand, we determine the same measures for the structured models that were theoretically derived (see Figures 2, 3, and 4). If the measures for the structured models turn out to be better than those for the unstructured models, this would indicate that the theoretic assignment of the indicators (items) to the different subscales is empirically supported (indication of nomological validity). At the same time, this would also be a hint that the individual subscales are represented by homogenous items (indication of internal consistency). Further, we expect that the fit measures for the structured models reach values that are deemed acceptable.

For conducting CFA, alternative estimation methods can be applied. Maximum Likelihood estimation, the most common method, requires the manifest variables to follow a multivariate normal distribution (Backhaus et al. 2006, 370). A necessary –though not sufficient– condition for multivariate normal distribution are univariate normal distributions of the individual variables. As our data does not meet this requirement (Kolmogorov-Smirnov-Test), we use the Asymptotically Distribution-Free method (ADF), which does not require normal distribution

(Backhaus et al. 2006, 370-371). However, this method requires relatively large sample sizes (*n*), which are determined depending on the number of manifest variables (*p*) according to the formula $n \geq 0.5p(p + 1)$ (*Gold et al.*, 2002, 12). This requirement is met for all three sections of our questionnaire.[6]

Assessment of the Overall Model

In a third step, we combine the three individual constructs' structured models to form an overall model (*fully structured overall model*, see Figure 5). This step particularly serves to allow assessments regarding discriminant validity of the measured constructs. To challenge our contention that the items of the three first sections of the questionnaire measure distinct constructs, we fist specified a comparative model, in which all items load on one and the same overall construct (*unstructured overall model*, see Figure 6). In a second comparative model, the items are structured according to the subscales described above, with the subscales loading all on one and the same overall construct (*semi-structured overall model with subscales*, see Figure 7).

[Insert Figures 5, 6, and 7 about here]

We expect that this second comparative model will achieve better fit measures than the unstructured comparative model, because this would lend empirical support to the theoretically derived structuring of the items and the subscales (indication of internal consistency of the subscales). We will finally compare the two comparative models with the fully structured overall model. The fit measures of the fully structured overall model should be better than the fit measures of both comparative models, if the items do indeed measure three distinct constructs

[6] The first section of the questionnaire (error handling) contains 12 items, thus requires 78 subjects, the second section (error climate) contains 15 items, thus requires 120 subjects, and the third section contains 13 items, thus requires 91 subjects. With our sample of 284 subjects all of these minimum numbers are clearly exceeded.

(indication of discriminant validity) and if they are homogenous within each construct (indication of internal consistency).

For the overall models, the ADF estimation method is not applicable, because our sample size is not big enough.[7] As the requirement of multiple normal distribution is not met either, we use the Scale-Free Least Squares method (SLS), which does not require multivariate normal distribution and which also works with small sample sizes. Due to the different estimation method, the fit measures are not comparable with those obtained for the individual models and subscales that were computed with the ADF method. However, this does not cause any problem, as only the relative comparison among the three differently structured models is relevant here. For assessing the three models, we compute the fit indices CMIN/DF, GFI, AGFI, and NFI.

As a further criterion to assess discriminant validity, we calculate the Fornell-Larcker-criterion for the three constructs, which is considered to be a particularly strict criterion (Homburg and Giering 1998, 126). It requires a construct's extracted variance to be bigger than the squared correlation of this construct with another construct (Fornell and Larcker 1981, 45-46).

RESULTS

Specific Measures of Reliability

First we calculated the indicator reliabilities of all items with regard to the corresponding subscales, as well as the construct reliabilities of the subscales for each section of the questionnaire.[8] The results are summarized in Table 2.

---

7 According to the formula provided by *Gold et al.* (2002, 12; see also above in the text), at least $n = 820$ subjects would be required for $p = 40$ items.

8 In section 1 of the questionnaire, we eliminated item 10 prior to the analyses, as it was the only item that was coded in reverse direction. Due to the skewness of the distribution, the inclusion of this item could bias the results. Another indication that supports this worry is a low item-to-total correlation for this item.

[Insert Table 2 about here]

The indicator reliabilities of all error handling items are below the limit value of 0.5. While the construct reliability of the subscale "oral feedback" is 0.62 and thus acceptable, the reliability of the subscale "written feedback" is only 0.52, i.e. below the lower limit of 0.6. The Cronbach Alpha values are also quite low (0.56 and 0.39, respectively).[9] However, the construct reliability of the overall construct "error handling" is 0.67 and thus acceptable (Cronbach Alpha = 0.56).

For organizational error climate, all indicator reliabilities are above the limit value. The subscales' construct reliabilities are between 0.83 and 0.93, which is clearly above the minimum value. Likewise, the reliability of the overall construct "error climate" is 0.97, which is an excellent value. The Cronbach Alpha values are also good (between 0.73 and 0.90 for the subscales and 0.86 for the overall construct).

For individual error orientation, indicator reliabilities range between 0.27 and 0.77, with six of the 13 items being below the limit value. However, the construct reliabilities of the subscales are all above the limit value (values between 0.65 and 0.85). The reliability of the overall construct "error orientation" is 0.90, which is a very good value. Cronbach Alpha values for the subscales range between 0.69 and 0.83; the value for the overall construct is 0.86.

Overall, all subscales except for "written feedback" of the error handling construct reach the required limit values. In most cases, the obtained values are clearly above the required limits. The items and subscales of both the organizational error climate construct and the individual error orientation construct reached good values, whereas some items and one subscale of the error handling construct obtained poorer values.

---

9 Cronbach Alpha values above 0.8 are deemed good (Bortz and Döring 2002, 196). However, for explorative investigations, in practice also far worse values are accepted (Schnell et al. 2005, 153).

Assessment of Individual Constructs' Models

The fit indices for the unstructured and structured individual models are summarized in Table 3.

[Insert Table 3 about here]

For all three constructs, the structural models reach better fit indices than the unstructured models, without any exception. These empirical findings support the theoretic assignment of the indicators to the different subscales in the structured models and thus are indications of nomological validity and internal consistency of the constructs and the measurement instruments.

Besides this comparative analysis, the calculated fit measures can also be interpreted in absolute terms by confronting them with experience-based values described in the literature. As upper limits for the CMIN/DF, values between 2.0 and 3.0 are usually used (Homburg and Giering 1998, 130; Homburg and Pflesser 2000, 430; Hair et al. 1998, 658). For all of our structured models, CMIN/DF is smaller than 3.0 and thus lies within an acceptable range. For the RMSEA, values smaller than 0.08 are considered acceptable (Browne and Cudeck 1993, 144; Hair et al. 1998, 656; Schermelleh-Engel et al. 2003, 36). For our structured models, RMSEA-values are below this limit. The GFI should be bigger than 0.9 (Schumacker and Lomax 1996, 121) and the AGFI bigger than 0.85 (Schermelleh-Engel et al. 2003, 43). For the structured models of error handling and error orientation, these minimum values are exceeded. Only for the structured model of error climate the values are below the limits. The suggested minimum values of the incremental fit indices are 0.9 for the NFI (Schumacker and Lomax 1996, 121; Hair et al. 1998, 657) and 0.95 for the TLI (Hu and Bentler 1999, 27; Schermelleh-Engel et al. 2003, 41-42; Marsh and Grayson 1995, 192). The values for all of our models are clearly below these limits.

To sum up, the structured models of error handling and error orientation reached the required values for all stand alone fit indices, whereas the incremental fit indices fell below the limit values. For the structured model of error climate, only the CMIN/DF and the RMSEA reached the minimum values. However, for all structured models all fit indices are better than for the unstructured models, without any exception.

Assessment of the Overall Model

The fit indices calculated for the fully structured overall model with three layers and for the two comparative models are summarized in Table 4.

[Insert Table 4 about here]

As expected, the fit measures improved with more detailed structuring of the models. Also in absolute terms, all fit indices for the fully structured overall model and even for the semi-structured model with subscales exceed the required limit values. These empirical findings for the fully structured model support the theoretically derived relationships that were specified within this model. In particular, this indicates that the error-handling-items, the error-climate-items, and the error-orientation-items measure three distinct constructs (indication of discriminant validity). It also indicates homogeneity of the three constructs (indication of internal consistency).

For further assessing discriminant validity, Table 5 provides the correlations among the three constructs, as well as the extracted variance for each construct, which are required for determining the Fornell-Larcker-criterion.

[Insert Table 5 about here]

For all three constructs, the extracted variance exceeds the squared correlations between the respective construct and the two other constructs. Thus, the Fornell-Larcker-criterion of discriminant validity is fulfilled for all three constructs.

DISCUSSION

Overall, both the specific reliability measures and the results of the CFA provide evidence for nomological and discriminant validity and for (internal-consistency-)reliability of our research instrument and the measured constructs, respectively. However, not all of the calculated measures for assessing validity and reliability reached the recommended limit values. Nevertheless, having in mind the early stage in developing a research instrument for measuring auditors' error handling behavior and related constructs, the results are more than satisfactory.

This is particularly true for the *error-handling*-construct and the corresponding items of the questionnaire, which were originally developed. However, for this construct specific reliability measures are still in need of improvement. For a considerable improvement, the items of the subscales would need to be more homogenous. Thus, for increasing overall reliability, further items that are as homogenous as possible could be added to the subscales. However, with such a strategy there is a danger of not sufficiently representing the different facets or dimensions of the (error-handling-)construct by the items. Therefore, it cannot be sufficient to add further homogenous items that just measure the existing subscales "oral feedback" and "written feedback" on a general level.

A more promising strategy should be to further differentiate these subscales according to theoretic considerations. Items that capture specific dimensions of the existing subscales would need to be developed. Following this idea, a further differentiation of the subscale "written feedback" seems to be possible. Within this subscale, some items capture aspects of the auditor's confrontation with the auditee, whereas others refer to a rather pedantic behavior like a complete listing of all detected errors, for example. A theoretic revision of the measurement model could explicitly take into consideration these dimensions, which in turn should be operationalized by means of appropriate additional items. As a starting point, however, the present model specification and its operationalization seem to be suitable. This conclusion is supported by the –altogether– good fit measures of the model.

The sections of the questionnaire that measure *error climate* and *error orientation* are based on questionnaires that have already been tested and applied in psychological research, which we adapted to an auditing context. For both constructs, the specific reliability measures yielded excellent results. Only the fit measures indicated that particularly for the error-climate-construct even better results might be achievable through a revision of the model.

Finally, the results provide substantial support for discriminant validity. That is, they provide strong indication that the first three sections of our questionnaire do indeed measure three distinct constructs (error handling, error climate, and error orientation).

As a limitation that might restrict the generalizability of the findings in this study, in principle these findings only apply to the investigated sample. For instance, most public sector auditors in our sample came from local governmental audit departments, so it remains open whether public sector auditors from other administrative levels or institutions would response differently from our subjects. Furthermore, subjects were not strictly selected by chance. Rather, data from an "arising" sample was used. However, we do not expect the measured variables to be biased within the sample, i.e. they should not be connected systematically with the probability of a potential subject to enter into our sample. So, we do not expect that the way of subject selection negatively affected the study's results.

Overall, the empirical support for the questionnaire's validity and reliability obtained in this study encourages the questionnaire's use in future research studies on error handling, error climate, and error orientation in auditing. Future research could investigate how the constructs of error handling, error climate, and error orientation are related to each other. Specifically, it could be studied whether the organizational error climate and the auditor’s individual error orientation influence the auditor’s error handling behavior, i.e. the auditor’s way of dealing with the auditee’s errors. Other potential determinants of error handling behavior such as the type and purpose of the audit, the error type, as well as further personal and situational determinants should also be included in a study of this type. Future research should also investigate which consequences error handling behavior may have for both the auditor and the auditee, including consequences for the auditor’s and the auditee’s organizations.

Knowledge about these determinants and consequences should help to derive recommendations for the auditor’s specific behavior concerning the auditee’s errors and for adequate organizational conditions. Moreover, such knowledge could be useful for determining which error handling strategy is appropriate for achieving desired consequences in specific situations. Finally, findings with regard to the costs of alternative error handling strategies might help to decide which strategy allows to achieve specific consequences at minimum audit costs. Thus, future research in this area seems to be important, as findings are likely to be relevant for both audit effectiveness and audit efficiency.

REFERENCES

Allen, R. D. 1996. Managing internal audit conflicts. *Internal Auditor* (August): 58-61.

Backhaus, K., B. Erichson, W. Plinke, and R. Weiber. 2006. Multivariate Analysemethoden: Eine anwendungsorientierte Einführung. 6th edition. Berlin and Heidelberg.

Baetge, J. 1970. Möglichkeiten der Objektivierung des Jahreserfolges. Düsseldorf.

Bagozzi, R. 1979. The role of measurement in theory construction and hypothesis testing: toward a holistic model. In *Conceptual and Theoretical Developments in Marketing*, edited by O. Ferell, S. Brown, and C. Lamb, 15-32. Chicago.

———, and L. Phillips. 1982. Representing and testing organizational theories: A holistic construal. *Administrative Science Quarterly*: 459-489.

———, and Y. Yi. 1988. On evaluation of structural equation models. *Journal of the Academy of Marketing Science* (Spring): 74-94.

Baker, F. W. 1978. The team approach to auditing. *Internal Auditor* (June): 58-62.

Bearden, W. O., S. Sharma, and J. E. Teel. 1982. Sample size effects on chi square and other statistics used in evaluating causal models. *Journal of Marketing Research*: 425-430.

Bollen, K. A. 1989. Structural equations with latent variables. New York et al.

Bortz, J., and N. Döring. 2002. Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. 3rd edition. Berlin et al.

Browne, M. W., and R. Cudeck. 1993. Alternative ways of assessing model fit. In *Testing Structural Equation Models*, edited by K. A. Bollen, and J. S. Long, 136-162. Newbury Park et al.

Buchner, R. 1997. Wirtschaftliches Prüfungswesen. Munich.

Burke, K. G., and S. E. Kovar. 1999. The power of persuasion: Negotiation strategies for internal auditors. *Internal Auditing* (March/April): 11-15.

Campfield, W. L. 1983. A look at responses to audit findings. *Internal Auditor* (October): 49-52.

Chadwick, W. E. 2002. Oh No! The Auditor is here! *Internal Auditor* (April): 52-55.

Crockett, J. R. 1992. Conducting the post-audit conference. Altamonte Springs.

Cronbach, L. J. 1951. Coefficient alpha and the internal structure of tests. *Pychonometrica*: 297-334.

Cudeck, R., and M. W. Browne. 1983. Cross-validation of covariance structures. *Multivariate Behavioral Research*: 147-167

Didis, S. K. 1997. Communicating audit results. *Internal Auditor* (October): 36-38.

Edmondson, A. C. 1996. Learning from mistakes is easier said than done: Group and organizational influences on the detection and correction of human error. *Journal of Applied Behavioral Science*: 5-28.

Egner, H. 1980. Betriebswirtschaftliche Prüfungslehre. Berlin and New York.

Fornell, C., and D. F. Larcker. 1981. Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research* (February): 39-50.

Frese, M. 1995. Error management in training: Conceptual and empirical results. In *Organizational Learning and Technological Chance*, edited by C. Zucchermaglio, S. Bagnara, and S. U. Stucky, 112-124. Berlin.

———, and C. van Dyck. 1996. Error management: Learning from errors and organizational design. Working Paper, University of Amsterdam.

Göbel, S., and M. Frese. 1999. Persönlichkeit, Strategien und Erfolg bei Kleinunternehmern. In *Unternehmerisch erfolgreiches Handeln*, edited by K. Moser, B. Batinic, and J. Zempel, 93-113. Göttingen.

Gold, M. S., P. M. Bentler, and K. H. Kim. 2002. A comparison of maximum-likelihood and asymtotically distribution-free methods of treating incomplete non-normal data. Working paper, University of California. Available at: http://repositories.cdlib.org/uclastat/papers/2002010124. Download: May 15th, 2006.

Haakenson, R. 1976. The art of listening. *Internal Auditor* (August): 32-40.

Hair, J. F. Jr., R. E. Anderson, R. L. Tatham, and W. C. Black. 1998. Multivariate data analysis. 5th edition. New Jersey.

Holman, R. 1981. Communication: An essential element of internal auditing. *Internal Auditor* (December): 39-44.

Homburg, C., and H. Baumgartner. 1998. Beurteilung von Kausalmodellen: Bestandsaufnahme und Anwendungsempfehlungen. In *Die Kausalanalyse: Ein Instrument der empirischen betriebswirtschaftlichen Forschung*, edited by L. Hildebrandt, and C. Homburg, 343-369. Stuttgart.

———, and A. Giering. 1998. Konzeptualisierung und Operationalisierung komplexer Konstrukte: Ein Leitfaden für die Marketingforschung. In *Die Kausalanalyse: Ein Instrument der empirischen betriebswirtschaftlichen Forschung*, edited by L. Hildebrandt, and C. Homburg, 111-146. Stuttgart.

———, and C. Pflesser. 2000. Konfirmatorische Faktorenanalyse. In *Marktforschung: Methoden, Anwendungen, Praxisbeispiele*, edited by A. Herrmann, and C. Homburg, 413-437. 2nd edition. Wiesbaden.

Hu, L., and P. M. Bentler. 1999. Cutoff criteria for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling*: 1-55.

Institute of Internal Auditors. 1972. Behavioral patterns in internal audit relationships. Research Committee Report 17. New York.

Jeffords, R., G. Thibadoux, and M. Scheidt. 2001. Understanding Social Styles for Internal Auditors. *Internal Auditing* (September / October): 17-24.

Keating, G. 1995. The art of the follow-up. *Internal Auditor* (April): 59-61.

Kimbrough, R. B. 1997. Facilitating trust. *Internal Auditor* (August): 64-66.

Kinney, W. 2000. Research on nature, characteristics and causes of accounting errors: The need for a multi-method approach. Discussant comments, *Journal of Accounting Literature*: 93-101.

Lindner, M. 2003. Interne Revision: Best Practice zur Weiterentwicklung der Internen Revision. In *Handbuch Management: Mit Best Practice zum Managementerfolg*, edited by C. Jahns, and G. Heim, 405-417. Stuttgart.

Lynch D. H., and S. Golen. 2003. 10 steps to writing documents. *Internal Auditor* (February): 53-57.

Maniak, A. J. 1990. Report writing for internal auditors. Chicago.

Marsh, H. W., and D. Grayson. 1995. Latent variable models of multitrait-mulitmethod data. In *Structural equation modeling: Concepts, issues, and applications*, edited by H. Hoyle, 177-198. Thousand Oaks et al.

Mautz, R. K., and H. A. Sharaf. 1961. The Philosophy of Auditing. Sarasota.

Nunnally, J. C., and I. H. Bernstein. 1994. Psychometric theory. 3rd edition. New York et al.

Richter, M. 1999. Der Prüfer als Veränderer ("change agent") – eine abwegige Forderung aus dem Elfenbeinturm der Wissenschaft oder ein notwendiges Leitbild für Prüfer und Prüfungsorgane in der öffentlichen Verwaltung? In *Stand und Perspektiven der Öffentlichen Betriebswirtschaftslehre*, edited by D. Bräuning, and D. Greiling, 351-360. Berlin.

———. 2000. Marketing der Prüfungsfunktion und der Prüfungsergebnisse – eine Notwendigkeit für öffentliche Prüfer. In *Neues öffentliches Rechnungswesen: Stand und Perspektiven*, edited by D. Budäus, W. Küpper, and L. Streitferdt, 655-673. Wiesbaden.

Ridley, J. 2003. Hat trick. *Internal Auditing & Business Risk* (August): 34-36.

Rybowiak, V., H. Garst, M. Frese, and B. Batinic. 1999. Error orientation questionnaire (EOQ): Reliability, validity, and different language equivalence. *Journal of Organizational Behavior*: 527-547.

Schermelleh-Engel, K., H. Moosbrugger, and H. Müller. 2003. Evaluating the fit of structural equation models: Test of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*: 23-74. Available at: http://www.mpr-online.de. Download: April 24th, 2006.

Schnell, R., P. B. Hill, and E. Esser. 2005. Methoden der empirischen Sozialforschung. 7th edition. Munich.

Schulte, E. B. 1970. Quantitative Methoden der Urteilsgewinnung bei Unternehmensprüfungen. Düsseldorf.

Schumacker, R. E., and R. G. Lomax. 1996. A beginner's guide to structural equation modeling. Mahwah.

Schwartz, B. M. 1999. Documenting audit findings. *Internal Auditor* (April): 48-49.

Sheffield, M. 1982. Real-world auditing. *Internal Auditor* (October): 32-33.

Sitkin, S. B. 1992. Learning through failure: The strategy of small losses. *Research in Organizational Behavior*: 231-266.

Sobel, M. E., and G. W. Bohrnstedt. 1985. The Use of null models in evaluating the fit of covariance structure models. In *Sociological Methodology*, edited by N. B. Tuma, 152-178. San Francisco.

Stier, W. 1999. Empirische Forschungsmethoden. 2nd edition. Berlin et al.

Van Dyck, C. 2000. Putting errors to good use: Error management culture in organizations. Dissertation, University of Amsterdam.

Von Wysocki, K. 1988. Grundlagen des betriebswirtschaftlichen Prüfungswesens. Munich.

Wanik, O. 1985. Prüfung des Internen Kontrollsystems. In *Handbuch der Abschlußprüfung: Ein systematischer Leitfaden aus dem Handwörterbuch der Revision*, edited by A. G. Coenenberg, and K. von Wysocki, 848-861. Stuttgart.

Wells, J. T. 2003. Rules for the written report. *Journal of Accountancy* (December): 75-78.

Whitham, R. B. 1996. Writing strategies for effective internal audit reports. *Internal Auditing* (Winter): 60-62.

Williams, M. 2003. The audit profession: Center stage. *Internal Auditor* (April): 50-57.

Zapf, D., M. Frese, and F. C. Brodbeck. 1999. Fehler und Fehlermanagement. In *Arbeits- und Organisationspsychologie: Betriebsbezogene Zielbereiche*, edited by D. Frey, C. Graf Hoyos, and D. Stahlberg, 398-411. Weinheim.

APPENDIX

FINAL QUESTIONNAIRE FOR PUBLIC SECTOR AUDITORS[10]

Section 1

In this section of the questionnaire, we would like to find out how you behave when you have detected an error of an employee or the audited entity during an audit. Please consider exclusively **unintentional mistakes**; behavior concerning malicious action shall not be investigated. Please indicate how often during an audit you behave as indicated in the following points! There are no "right" or "wrong" answers.

	How often does it happen, that you	**never**	**rarely**	**sometimes**	**often**	**regularly**
1	conduct additional audit procedures, without informing the auditee about the goal of the audit?	❑	❑	❑	❑	❑
2	do not give any feedback to the auditee about the audit findings, in order to prevent him from hindering the continuing audit?	❑	❑	❑	❑	❑
3	include suggestions for improvement/measures into the report that you have not yet conferred about with the auditee?	❑	❑	❑	❑	❑
4	do not give an error feedback to the one how caused the error, because he is not well versed with the subject matter?	❑	❑	❑	❑	❑
5	give an error feedback directly to the superior of the one responsible for the error?	❑	❑	❑	❑	❑
6	talk to the auditee after detecting a deviation, but without telling him the scale of the error?	❑	❑	❑	❑	❑
7	behave consciously distant in front of the auditee?	❑	❑	❑	❑	❑
8	mention errors and their responsibilities clearly in the report, although the errors were already eliminated during the audit?	❑	❑	❑	❑	❑
9	propose organizational measures in the report, which the auditee has rejected in a conversation?	❑	❑	❑	❑	❑
10	do without listing smaller errors in the report that have already been eliminated, and instead note that "minor discrepancies could be clarified with the auditee"?	❑	❑	❑	❑	❑
11	do **not** mention the auditee's argumentation that differs from your finding?	❑	❑	❑	❑	❑
12	also include minor errors in the report?	❑	❑	❑	❑	❑

[10] This is the translated version (from German to English). The questionnaires for external and internal auditors only differ slightly in some wordings and demographic data requests from the public-sector-questionnaire. They were omitted here.

Section 2

In this section of the questionnaire, we would like to know which climate prevails in your administration. Please indicate to what extent every statement applies **to your administration in general**. Please think of **own errors** that have occurred when you were doing your job. Here as well there are no “right” or “wrong” answers.

	To what extent does this statement apply for your organization in general?	**applies completely**	**rather applies**	**neither nor**	**rather does not apply**	**does not apply at all**
1	In our organization there is agreement on that one can learn a lot from ones errors when doing ones work.	❑	❑	❑	❑	❑
2	If one cannot continue to work after an error, one can direct oneself at another colleague at any time.	❑	❑	❑	❑	❑
3	In this organization we know that it is appropriate to analyze errors.	❑	❑	❑	❑	❑
4	If an error occurred, it is always carefully analyzed.	❑	❑	❑	❑	❑
5	All errors are openly discussed.	❑	❑	❑	❑	❑
6	When an error occurred, it is immediately corrected.	❑	❑	❑	❑	❑
7	If one has made a mistake, one tries to analyze the error’s cause.	❑	❑	❑	❑	❑
8	If something went wrong, one devotes time to think it through.	❑	❑	❑	❑	❑
9	In general, the colleagues mutually warn each other if errors may occur.	❑	❑	❑	❑	❑
10	If one has made a mistake, one receives the full support by the colleagues.	❑	❑	❑	❑	❑
11	Errors are discussed among the colleagues.	❑	❑	❑	❑	❑
12	Quick and appropriate correction of errors is of highest priority.	❑	❑	❑	❑	❑
13	Because errors and their solutions provide important information for our work, we discuss them.	❑	❑	❑	❑	❑
14	When someone makes a mistake, he / she tells the others about it, so that they do not repeat the same mistake.	❑	❑	❑	❑	❑
15	If one cannot correct an error oneself, one directs oneself at a colleague.	❑	❑	❑	❑	❑

Section 3

The individual attitude towards errors can be quite different among different persons. In this section of the questionnaire we would like to know what an error means **for you personally**. Please think of **own errors** during daily work that have **occurred to you**. Please have in mind that your personal opinion is requested. There are no "right" or "wrong" answers.

	To what extent does this statement, regarding your own errors, apply for yourself?	**applies completely**	**rather applies**	**neither nor**	**rather does not apply**	**does not apply at all**
1	When I do not know how to proceed after an error, I can rely a 100% on my colleagues.	❑	❑	❑	❑	❑
2	From own errors I have already learned a lot for the doing of my work.	❑	❑	❑	❑	❑
3	When I have made a mistake in my work then I tell my colleagues about it, so that they do not make the same mistake.	❑	❑	❑	❑	❑
4	When I cannot rectify an error on my own, then I can direct myself at a colleague at any time.	❑	❑	❑	❑	❑
5	After an error I intensively think about how it is to be rectified.	❑	❑	❑	❑	❑
6	When an error is correctable I always know what I have to do.	❑	❑	❑	❑	❑
7	When I make a mistake in may work, then I always rectify it immediately.	❑	❑	❑	❑	❑
8	I often think about how I could have avoided an error.	❑	❑	❑	❑	❑
9	My errors show me what I must do better.	❑	❑	❑	❑	❑
10	When something at my work has not worked well, then I devote time to think about it.	❑	❑	❑	❑	❑
11	When I have made a mistake this is an important source of information for the doing of my job.	❑	❑	❑	❑	❑
12	When I have made a mistake I carefully analyze it.	❑	❑	❑	❑	❑
13	After I have made a mistake I think how it could happen.	❑	❑	❑	❑	❑

Finally, for the statistical analysis of the questionnaire, we need some personal details. Please tick the correct response-alternative for each question!

How many years of professional experience in public auditing do you have?

- ❑ less than 1 year
- ❑ 1-2 years
- ❑ 2-5 years
- ❑ more than 5 years

Do you predominantly audit the same departments:

- ❑ several times a year,
- ❑ one time a year,
- ❑ every 2-3 years, or
- ❑ less than every 3 years?

How satisfied are you with your work?

- ❑ very satisfied
- ❑ rather satisfied
- ❑ neither satisfied nor dissatisfied
- ❑ rather dissatisfied
- ❑ very dissatisfied

How many auditors work at your audit institution?

....

How old are you?

- ❑ under 30
- ❑ 31-40
- ❑ 41-50
- ❑ over 50

Are you:

- ❑ feminine,
- ❑ masculine?

How is the numeric relationship between auditors and employees of the audited administration?

- ❑ up to 1:500
- ❑ 1:501 to 1:1.000
- ❑ 1:1001 to 1:3.000
- ❑ above 1:3.000

In which audit institution do you work?

- ❑ supreme audit court
- ❑ local governmental audit department
- ❑ internal audit department of a public administration
- ❑ other:................................

We thank you sincerely for your participation in the study!

Please send the completed questionnaire to:

[…]

If you are interested in the study's results, you are invited to send and e-mail to the following address: […] After the analysis of the results we will be happy to send you a short report.

TABLE 1
Demographic Data of the Sample[11]

	Public Sector Auditors (Percent)	Internal Auditors (Percent)	External Auditors (Percent)
Professional Experience			
Less Than 1 Year	0	8	0
1-2 Years	2	23	0
2-5 Years	25	29	12
More Than 5 Years	72	40	87
Gender			
Women	75	37	13
Men	24	62	86
Age			
Under 30	4	25	2
31-40	25	44	49
41-50	44	21	18
Over 50	25	8	28

[11] Due to missing values, the values do not sum up to a hundred percent.

TABLE 2
Indicator and Construct Reliabilities for Sections 1 to 3 of the Questionnaire

		Indicator Reliability	Construct Reliability	Cronbach Alpha
Error handling			**0.67**	**0.56**
Oral Feedback	Item 1	0.32	**0.62**	**0.56**
	Item 2	0.41		
	Item 4	0.26		
	Item 5	0.08		
	Item 6	0.26		
	Item 7	0.06		
Written Feedback	Item 3	0.10	**0.52**	**0.39**
	Item 8	0.19		
	Item 9	0.25		
	Item 11	0.18		
	Item 12	0.19		
Error Climate			**0.97**	**0.94**
Help	Item 2	0.53	**0.86**	**0.81**
	Item 10	0.79		
	Item 15	0.71		
Analysis	Item 1	0.57	**0.92**	**0.90**
	Item 3	0.64		
	Item 4	0.83		
	Item 7	0.77		
	Item 8	0.72		
Communication	Item 5	0.67	**0.93**	**0.88**
	Item 9	0.66		
	Item 11	0.74		
	Item 13	0.85		
	Item 14	0.66		
Correction	Item 6	0.74	**0.83**	**0.73**
	Item 12	0.67		
Error Orientation			**0.90**	**0.86**
Communication	Item 1	0.77	**0.84**	**0.81**
	Item 3	0.40		
	Item 4	0.73		
Learning	Item 2	0.39	**0.73**	**0.69**
	Item 9	0.44		
	Item 11	0.59		
Analysis	Item 5	0.46	**0.85**	**0.83**
	Item 8	0.31		
	Item 10	0.60		
	Item 12	0.75		
	Item 13	0.55		
Competence	Item 6	0.27	**0.65**	**0.62**
	Item 7	0.72		

TABLE 3
Fit Indices for Individual Constructs' Models

		stand alone fit indices				incremental fit indices		
		Stand Alone Fit Indices				Incremental Fit Indices		
		Measures of Statistical Inference		Descriptive Measures				
				Without Consider-ation of DF	With Consider-ation of DF	Without Consider-ation of DF	With Consider-ation of DF	
		CMIN/DF	**RMSEA**	**GFI**	**AGFI**	**NFI**	**CFI**	**TLI**
Error Handling	Unstructured	2.454	0.072	0.933	0.900	0.414	0.505	0.381
	Structured	2.017	0.060	0.946	0.917	0.529	0.662	0.567
Error Climate	Unstructured	2.942	0.083	0.788	0.717	0.499	0.588	0.519
	Structured	2.665	0.077	0.816	0.744	0.567	0.662	0.588
Error Orientation	Unstructured	2.958	0.083	0.857	0.800	0.617	0.700	0.640
	Structured	2.065	0.061	0.907	0.861	0.749	0.847	0.804

TABLE 4
Fit Indices for Overall Models

		Stand Alone Fit Indices				Incremental Fit Indices		
		Measures of Statistical Inference		Descriptive Measures				
				Without Consider-ation of DF	With Consider-ation of DF	Without Consider-ation of DF	With Consider-ation of DF	
		CMIN/DF	**RMSEA**	**GFI**	**AGFI**	**NFI**	**CFI**	**TLI**
Overall Models	Unstructured (One Layer)	3.094	NA	0.901	0.889	0.867	NA	NA
	Semi-Structured with Subscales (Two Layers)	2.005	NA	0.936	0.928	0.915	NA	NA
	Fully Structured (three layers)	1.652	NA	0.948	0.941	0.930	NA	NA

TABLE 5
Fornell-Larcker-Criterion (Squared Correlations and Variance Extracted)

Squared Correlation Between	Error Handling	Error Climate	Error Orientation	< ?	**Variance Extracted**
Error Handling	---	0.002	0.034	**<**	0.507
Error Climate	0.002	---	0.513	**<**	0.894
Error Orientation	0.034	0.513	---	**<**	0.696

FIGURE 1
Conceptual Framework of Error Handling, Error Climate, and Error Orientation

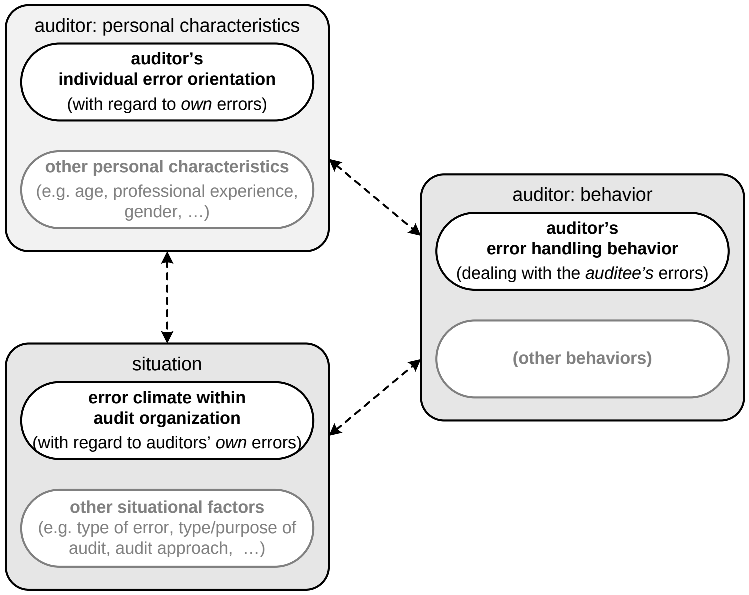

FIGURE 2
Measurement Model for Error Handling (Structured Model)

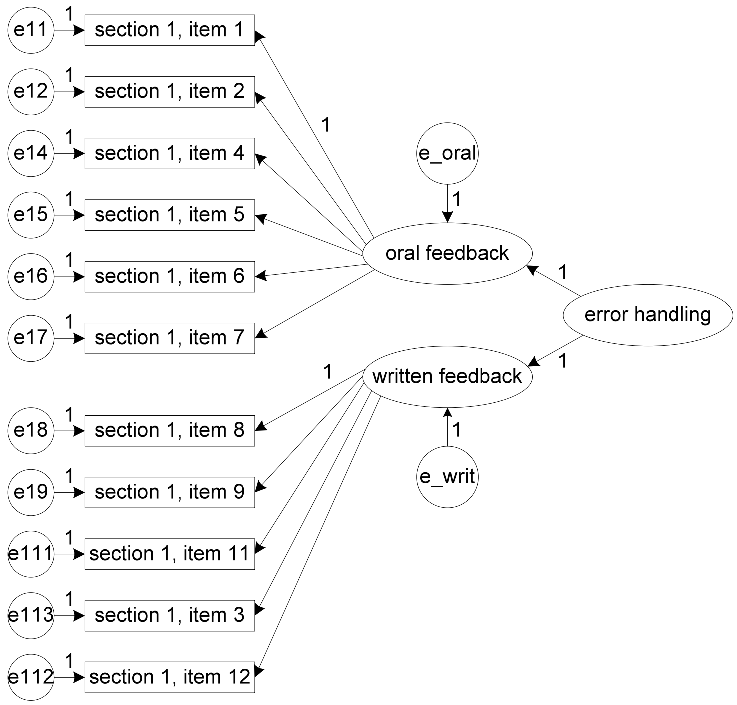

FIGURE 3
Measurement Model for Error Climate (Structured Model)

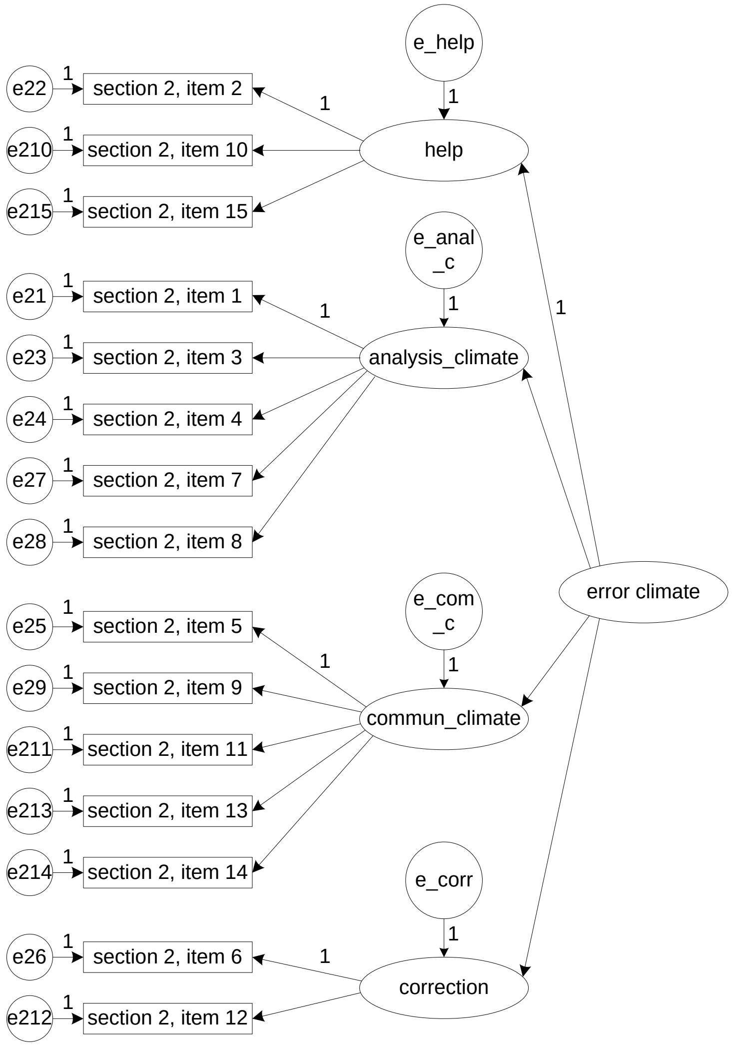

FIGURE 4
Measurement Model for Error Orientation (Structured Model)

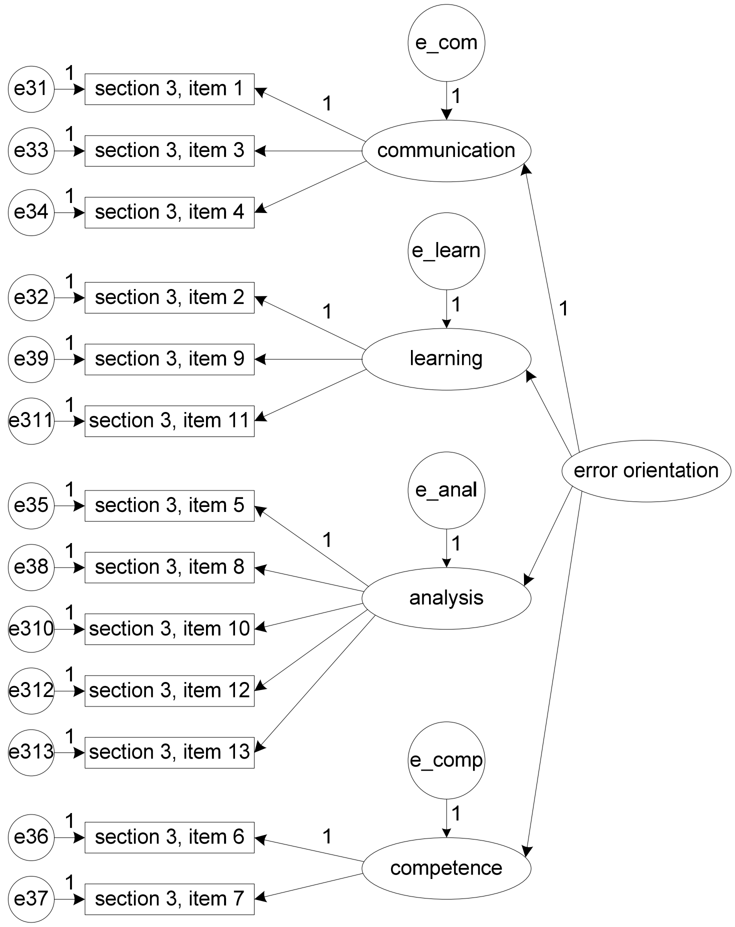

FIGURE 5
Fully Structured Overall Model (Three Layers)

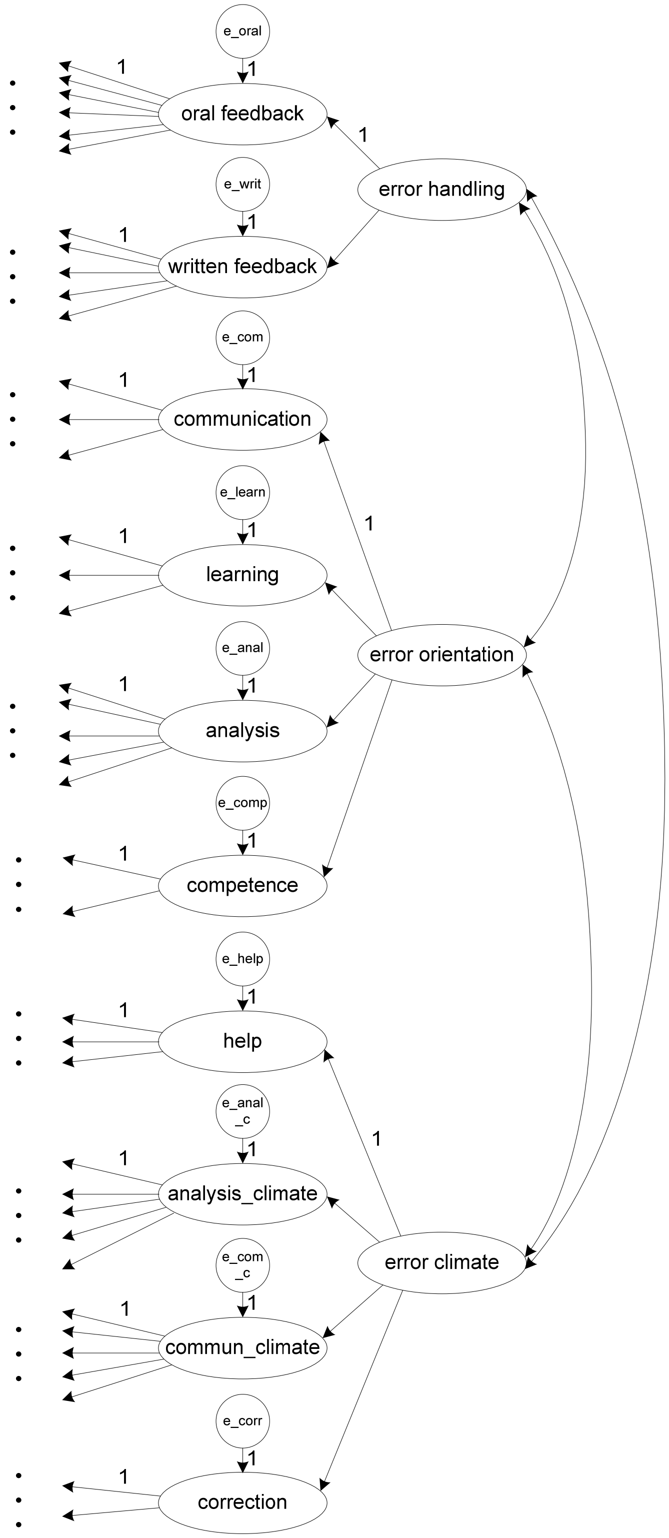

FIGURE 6
Unstructured Overall Model (One Layer)

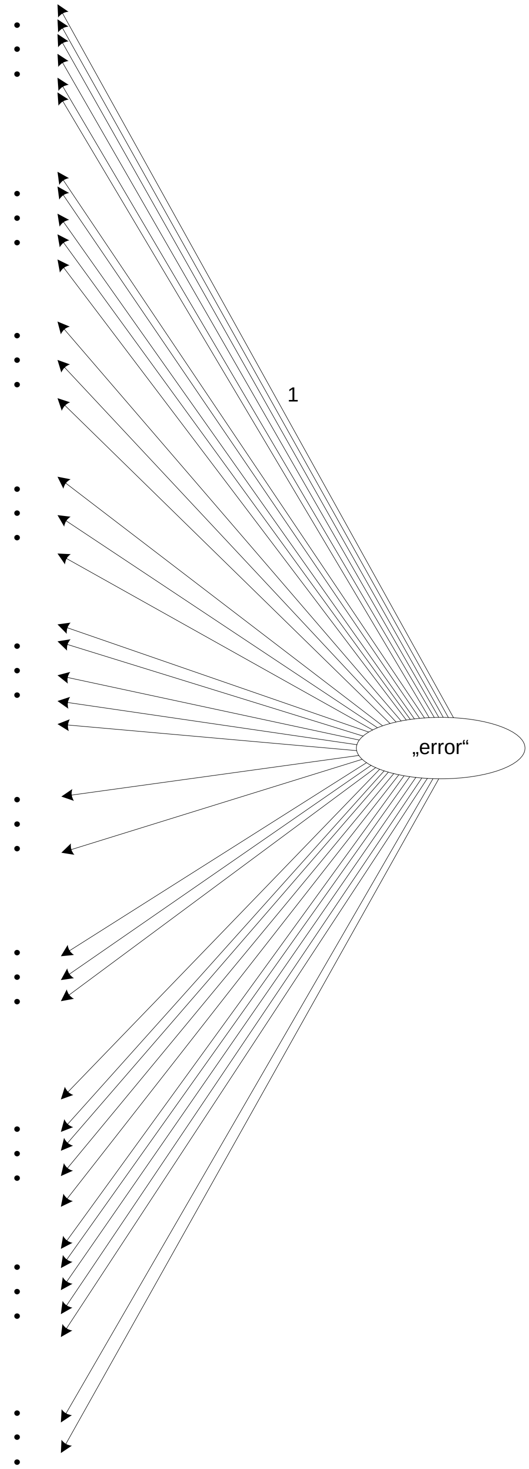

FIGURE 7

Semi-Structured Overall Model with Subscales (Two Layers)